DOCUMENT RESUME

ED 093 946 TM 003 764

AUTHOR	Peck, Robert F.; Borich, Gary
TITLE	Personality Measures That Predict Teaching Performance.
INSTITUTION	Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C. Researcher Training.; Office of Education (DHEW), Washington, D.C.
PUB DATE	Apr 74
CONTRACT	NE-C-00-3-0066; OEC-6-10-108
NOTE	27p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)
AVAILABLE FROM	R & D Center for Teacher Education, The University of Texas at Austin, Education Annex 3.203, Austin, Texas 78712 ($3.29)
EDRS PRICE	MF-$0.75 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS	Attitude Tests; *Individual Characteristics; *Performance; *Personality Tests; *Predictor Variables; Self Evaluation; Student Evaluation; Student Teachers; Teachers; *Teaching

ABSTRACT

Personality and attitude measures taken before training (of 70 primary grade student teachers and 71 teachers of grades 4-9) were compared, by multiple linear regression analyses, with instructor, self, and pupil evaluations of their teaching performance, at the end of training. Pupil evaluations were highly reliable; instructor and self-ratings, not very reliable. The personality variables explained a substantial part of the variance in teaching performance is measured by pupils, though not by instructors or by self. The predictive personality patterns were quite different in grades 1-3 and in grades 4-9. They are described. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Ses

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Robert Peck
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

Personality Measures That Predict Teaching Performance

Robert F. Peck and Gary Borich

Research and Development Center for Teacher Education

The University of Texas at Austin

ED 093946

Copyright 1974 by Robert F. Peck

This paper was written for presentation at the annual meetings of the American Educational Research Association, April 15-19, 1974, in Chicago, Illinois.

The research reported herein was supported in part by U. S. Office of Education Contract OE-6-10-108 and National Institute of Education Contract NE-C-00-3-0066, Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education or the National Institute of Education and no official endorsement by those offices should be inferred.

We are gratefully indebted to the administration and the many teachers in the Austin Independent School District who cooperated in this study; to the school children who took their part seriously and well, to our colleagues in the College of Education at The University of Texas at Austin for their thoughtful cooperation and time; and to Lee Cronbach for an extremely thoughtful, detailed critique of an earlier draft of this report which has substantially improved its accuracy and clarity.

TM 003 764

Personality Measures That Predict Teaching Performance

Robert F. Peck and Gary Borich

Research and Development Center for Teacher Education

The University of Texas at Austin

The instruments in the Comprehensive Personal Assessment System (COMPASS) battery, measuring attributes of personality, attitude, values and motivation, are normally subjected to an integrated, clinical interpretation for use in Personalized Assessment Feedback counseling, in the Personalized Teacher Education Program (Peck, 1970; Peck et al., 1972). In that application, the subscale scores on various instruments are never taken in isolation as predictors of behavior. Nonetheless, if the instruments do measure characteristics which are related to teaching behavior and teaching effectiveness, even an atomistic analysis should show some significant relationships of the subscale scores to measures of teaching performance. The present study was therefore designed to test the hypothesis that a significant number of the subscales on three quantitatively scorable instruments in the COMPASS battery will be significantly related to two kinds of measures of student teaching performance: ratings by the pupils of the student teacher; and ratings by the supervising teacher, by the university supervisor, and by the student herself.

The Sample

The sample consisted of 188 elementary student teachers. This was the total number of such students at The University of Texas at Austin who completed their pre-service teacher education in the Spring of 1972. Thus, it presumably could be taken as an unselected, representative sample of teacher-graduates from this institution. Complete data were ultimately obtained on 141 students: 70 who taught grades 1-3, and 71 who taught grades 4-9.

The Measures

The independent, predictor measures of attitude and personality consisted of the following scores, obtained at the time the sample enrolled in their first course in teacher education. In almost all cases, this was 8 to 20 months prior to the completion of student teaching.

Adjective Self-Description (ASD). Responses indicate the extent to which the respondent feels that each of 56 adjectives describes him. Seven factorially independent scales (Veldman and Parker, 1970) are measured:

1. Attitude. This scale corresponds to social warmth. A high scale score reflects a positive attitude or high social warmth. Representative items are "cheerful," "gentle," "good-natured."

2. Behavior. This scale corresponds to social abrasiveness or hostility. A high score indicates high hostility. Representative items are "obnoxious," "indifferent," "rude."

3. Efficiency. This scale corresponds to ego organization. A high score indicates high efficiency. Representative items are "efficient," "industrious," "organized."

4. Introversion. A high score on this scale indicates high introversion and low extroversion. Representative items are "quiet," "reserved," "shy."

5. Anxiety. A high score on this scale indicates high anxiety. Representative items are "anxious," "emotional," "moody."

6. Idealism. This scale corresponds to measures of individualism. A high score represents high idealism or individualism. Representative items are "complicated," "idealistic," "individualistic."

7. Attractiveness. A high score represents high attractiveness. Representative items are "charming," "good-looking," "sexy."

Self-Report Inventory (SRI). Responses indicate the extent to which the respondent feels that each of 48 statements describes him. Eight factorially distinct scales are measured (Bown and Veldman, 1967):

1. Self. A high score represents a positive attitude toward self. Sample item: "In almost every respect, I'm glad to be the person I am."

2. Others. A high score represents a positive attitude toward others. Sample item: "The way I get along with my friends is extremely important to me."

3. Children. A high score represents a positive attitude toward children. Sample item: "I'm very comfortable and happy when I am with children."

4. Authority. A high score represents a positive attitude toward persons in authority. Sample item: "I really enjoy getting to know people in positions of authority."

5. Work. A high score indicates a positive attitude toward work. Sample item: "Doing a good job in anything I undertake is very important to my sense of well-being."

6. Reality. A high score represents a positive attitude toward reality. Sample item: "I live in accordance with the idea that 'It is better to have loved and lost than never to have loved at all.'"

7. Parents. A high score represents a positive attitude toward parents. Sample item: "I am very happy with my present relationship with my parents."

8. Hope. A high score represents a positive attitude about the future. Sample item: "I really look forward to the time when I will be settled down to my life's work."

One-Word Sentence Completion (OWSC). Responses indicate the person's free associations to 62 items, selected to represent a broad array of developmental experiences and important intrapersonal and interpersonal stimuli, including a number drawn from the realm of teaching (Peck, 1960; Peck and Veldman, 1966; Veldman, 1971). Quantified, factorially distinct scores are obtained on seven scales:

1. Response Length. Average number of letters per response.

2. Repetitions. Number of times a response word is used in reply to two or more stems.

3. Populars. Responses given by more than 10% of the standardization sample (1718 students enrolled in the introductory Educational Psychology course at UTA, 1968-69).

4. Evasion. Includes failures to respond; the use of private, proper names; repetition of the key word in the stem; cryptic references; apparently deliberate ambiguity; and numeral responses.

5. Hostility. Responses express antagonism toward or devaluation of other people, or projection of blame for personal dissatisfaction.

6. Anxiety. Responses indicate apprehension, self-doubt, unusual fears or abnormal tension.

7. Depression. Responses indicate self-derogation, loneliness or depression.

The dependent criterion measures of teaching performance consisted of the following, collected at the end of student teaching:

Pupil Evaluations: SET II (Primary grades, 1-3)

Four factorially distinct scales were measured:

1. Stimulating, Interactive Rapport

2. Effective Teaching

3. Fosterance of Pupil Self-Esteem

4. Lack of Unreasonable Negativity

Pupil Evaluations: SET I (Upper grades, 4-9)

Five factorially distinct scales were measured:

1. Friendly and Cheerful

2. Knowledgeable and Poised

3. Lively and Interesting

4. Firm Control

5. Non-Directive Teaching Style

Teacher Educator Ratings. These were the mean scores across twelve items of the Teacher Evaluation Form: (1) Subject-Matter Competence; (2) Communication Skills; (3) Stimulating Imagination; (4) Responsible Independence; (5) Poise and Self-Confidence; (6) Attitude Toward Students; (7) Teaching Style; (8) Disciplinary Control; (9) Attitude Toward Supervision; (10) Reaction to Own Mistakes; (11) Alertness to Classroom Events; (12) Professional Commitment.

The ratings were independently performed by:

1. The student's Cooperating (or Supervising) Teacher

2. The student's University Supervisor

Self-Rating. This was the mean score on the Teacher Evaluation Form, responded to by the student teacher herself.

Reliability of the Criterion Measures

The reliability of the pupil evaluation forms, SET I and SET II (see Table 1) makes them substantial criteria for testing the predictive power of the personality measures (Veldman and Peck, 1964; Veldman, 1970; Haak, Kleiber, Peck, 1972).

The ratings by the teacher educators (and, less surprisingly, by the students themselves) turned out to be less reliable. Intercorrelation of the mean (overall) TEF ratings by the three sets of judges found the cooperating teacher and university supervisor correlation to be .60. These two sets of adult judges correlated only .39 and .28, respectively, with TEF self-ratings by the student teachers. When all eleven TEF scales were correlated across the three sets of judges, factor analyzed, and subjected to both varimax rotation and an image covariance analysis, the three different sets of judges were each represented by separate, unique factors. That is, there was very little agreement, even between cooperating teacher and university supervisor, when their detailed ratings of the students were examined.

One feature of such ratings is their skewness. Supervisors of student teachers are notorious for being "kind to everybody" in their grading and rating. To enhance students' chances to get jobs, and also to protect students' feelings, it would appear, cooperating teachers and university supervisors have long been known to skew their ratings heavily, using mostly the top end of any rating scale. (This is not peculiar to education; the same problem exists, to the same extreme degree, with individual fitness ratings in the armed forces.)

On a seven-point scale, the mean self-rating score in this sample was 5.3 (S.D. .63), considerably above the theoretical mid-score of 4.0. The ratings by university supervisors had a mean of 5.2 (S.D. .85), barely lower than the average of the self-ratings. Supervising teachers were the most generous of all, giving the student teachers an average score of 5.5 (S.D. .90); they rated extremely few student teachers even as low as "average," and rated most of them close to the top.

As for the relationships among these raters' perceptions, Table 2 shows the very low correlations found among the pupil, supervisor and self-evaluations. Quite clearly, student teachers' self-evaluations bore little, if any, relationship to evaluations either by pupils or by adult observers. Just as clearly, adult supervisors, whether school or campus based, did not agree too closely among themselves, and they agreed very little with pupils in evaluating student teacher performances.

Insofar as the agreement among pupils, as a large group of long-time observers, was more reliable than the agreement between the two adults who constituted the evaluators of any one student teacher, the SET scores would appear to be somewhat more stable criteria than the adult TEF ratings for validating any kind of predictor measures, including the personality and attitude measures in the COMPASS battery. The degree of reliability of the mean TEF scores, reflected by the cooperating teacher - university supervisor correlation of .60, was enough, however, to warrant testing the predictor measures against the TEF mean scores, as an additional estimate of the validity of the personality measures as predictors of teaching performance.

Method of Analysis

A multiple linear regression analysis was performed, comparing the 23 personality measures, as a set, with each of the 12 criterion measures of teaching behavior. Table 3 shows the Pearson correlations and, in the bottom line, the multiple correlation coefficient for each criterion.

The multiple R's were significantly large for three of the SET I criteria, and approached the .05 level of significance for the other criteria; but individual predictors' specific contributions to most of the criteria were not strongly defined by the simple correlations shown in Table 3, making uncertain the

description of the relationships between personality characteristics and teaching behavior. In order to determine which predictors might be making separate, significant contributions to each criterion an additional analysis was performed.

The results of this analysis are shown in Table 4. They address the question, what properties of the student teachers best account for the criterion performance? The simple Pearson r's, derived in the course of the multiple linear regression analysis, do not tell whether, nor how much, a predictor scale, such as ASD-Attitude, is a pure, uncontaminated measure of the properties of social warmth its definition states. Each Pearson r between a predictor variable and a criterion represents a mixture of the variance due to "social warmth," and the variance due to other properties of the teachers, including such properties as are measured by other scales in the battery that share some common degree of variance (i.e., significant correlations) with the criterion.

Consequently, a better estimate of the "true" properties of the teachers that account for the teaching performance criteria is derived by a part correlation analysis that shows how much of the variance on a criterion is uniquely predicted by a given predictor scale. Such an analysis identified the "pure" contribution of the property that is presumably designated by the name of a predictor scale ("pure," that is, within the set of predictor variables that have been measured). Each scale whose part correlation with a performance criterion reached statistical significance could thus be used to describe that property in the teachers that helped to account for their performance on the criterion measure. (Currently, there is debate about this approach. Its present use should therefore be viewed as a methodological experiment with an arguable method.

The Results section, below, draws on both kinds of analyses. For each criterion, the Pearson correlations and the multiple regression coefficient is given in Table 3, showing the degree to which the composite assessment battery predicted that criterion, in a purely empirical way. Thereafter, the results of the part correlation analysis are shown in Table 4, as a second kind of evidence about the relationships between teacher characteristics and teaching performance.

The percent of variance accounted for by each individual scale score of the ASD, SRI, and OWSC, for criterion behaviors measured by the TEF, SET I, and SET II, was determined for each predictor-criterion relationship, with all other predictors as covariates. The seven scales of the ASD, eight scales of the SRI, and seven scales of the OWSC comprised the predictors for this study, while the five factors of the SET I, four factors of the SET II and three forms of the TEF were the criterion measures. The percent of variance shown in the second-last line of Table 4 is the unique variance contributed by each predictor separately for each criterion. The main part of Table 4 shows the predictors that significantly ($p < .05$) increased the amount of criterion variance accounted for. Other variables that were significant predictors between the .10 and .06 levels of significance are also noted in Table 4, in parentheses, to indicate additional trends in the data. The effect of joint contributions of two or more predictors to criterion variance was ruled out by the covariance procedure.

The procedure employed in this "unique contribution" analysis was to construct a 22 predictor, multiple linear regression model as a "full" model and to successively drop out each predictor in turn to construct "restricted" models. The amount of variance accounted for was the difference between the squared

multiple R of the full model and the squared multiple R of the restricted model. The probability of this difference is given by:

$$F = \frac{R^2_{full} - R^2_{res} \;/\; (N - 22) - (N - 21)}{1 - R^2_{full} \;/\; (N - 22)}$$

for which df = 1/70, 1/71, or 1/141 for the SET I, SET II, and TEF, respectively. This test was repeated 22 times for each of 12 criteria for a total of 264 analyses.

Findings

Prediction of SET Scores from Personality Characteristics of the Student Teachers

The main hypothesis of the study was confirmed to a modest degree. As the bottom line of Table 3 shows, the multiple R's were sizable enough to reach or exceed the .05 level of significance for three of the five SET I criteria and for the Self-rating. The multiple R's for the other two SET I criteria and the four SET II criteria approached, but did not reach, the .05 level of significance. The ratings by adult supervisors were also predicted to a degree just short of the .05 level of significance. The data in Table 4 show that of the 23 personality characteristics measured by the three instruments, 19 accounted for a significant fraction of the variance in one or more of the criterion measures of teaching performance. The amount of variance thus "uniquely" explained ranged from a trivial 4% of the overall Supervisor rating and 5% of the overall Self-rating, to 34% of the SET I factor of Non-Directiveness; and the ratio of these figures to the total amount of variance "explained" by the total predictor battery ranged from 5/28 in the case of the overall Self-rating to 19/24 in the case of SET II Rapport and Lack of Negativity, and 34/43 in SET I Non-Directiveness.

(See last two lines of Table 4.) Thus, on most of the criteria, from two to six of the personality characteristics (not always the raw scores on the "same" predictor measures) accounted for a major part of the criterion variance "explained" by the total assessment battery.

Grades 1-3

Stimulating, Interactive Rapport is the first of the four SET II factors on which the primary children evaluated their student teachers. This aspect of teaching was predicted by the total set of personality scores, in a multiple linear regression analysis, just short of the .05 level of significance (multiple R = .49). Five personality attributes contributed substantially to this relationship. Teachers whose pupils rated high on Rapport were less likely than others in the sample to describe themselves as extremely warm and optimistic (ASD-Attitude) or as highly positive toward their parents (SRI). They showed, on the One-Word Sentence Completion, less hostility, a little more anxiety and somewhat more depressive feelings than student teachers with lower Rapport ratings. Interpreting this pattern just a little, it seems that young teachers who do not project an image of great optimism, who are working free of their dependency on parents, and who tend to blame themselves when things go wrong rather than get angry at other people, are more likely to generate a pleasant, lively, perhaps somewhat equalitarian rapport with primary grade children. By implication, they may be more conscientious than average and not so self-assured as to overwhelm children.

The pupil estimate of Effective Teaching showed a pattern very similar to that just described. Teachers rated high on this factor tended to be less extremely positive in expressing optimism and warmth (ASD-Attitude), less extremely positive in their attitude toward their parents (SRI), and somewhat more anxious (OWSC). In addition, they did not tend to respond with stereotypic

"Popular" answers on the OWSC. By implication, primary pupils find Rapport-stimulating properties an important part of good teaching, with the addition that they find "good teachers" thinking and talking in discriminating, non-stereotyped ways.

The multiple R between the predictor battery and this criterion was .46. The two simple r's that reached significance were for the ASD-Anxiety score and the OWSC-Anxiety score; both correlated positively with this criterion. The other correlations in Table 3 that corresponded to significant part-correlations in Table 4 showed the same directional relationship with the criterion.

Fosterance of Pupil Self-Esteem was related to: the teachers' extroversion (ASD-Introversion); self-liking (SRI-Self); no more than moderately positive attitudes toward their parents; a relatively brief, concise way of talking (OWSC-Response Length); a lack of stereotypy (OWSC-Populars); and a lack of hostility. Even young children, it appears, find that their growth in self-esteem is facilitated not only by receiving positive, pleasant feelings from the teacher but also by the teacher's own self-regard, her progress toward emotional emancipation from her parents, and her intellectual enterprise.

The multiple R of all predictor measures with this criterion was .45. The only disparity in the coefficients in Tables 3 and 4 was the simple r of .03 for OWSC-Populars, as compared with the part-correlation of -.27. Removing the criterion variance shared by this variable with all the other variables revealed a stronger unique contribution to the criterion than the simple correlations indicated. This was true of the "Populars" measure for many of the other criteria, as well.

Lack of Unreasonable Negativity was related to: a lack of strong emphasis on work (ASD); the same less-than-perfectly-positive attitude toward own parents, as above; conciseness of verbal expression (OWSC-Response Length);

a lack of stereotyped thinking (OWSC-Populars); less than average hostility (OWSC); and a somewhat higher incidence of depressed feelings (OWSC). This factor finds the pupils taking account of a good deal more than just the feeling tone the teacher displays, in judging the "reasonableness" of the teacher's moods. The children also connect this with the teacher's intellectual autonomy (if the low OWSC-Popular score may be interpreted this way). At this age, furthermore, the children find that teachers with a strong streak of the work-oriented "Protestant ethic" are not entirely reasonable or pleasant in their emotional reactions. This last point does not seem to reflect merely an aversion to effort on the children's part, when the rest of this correlational complex is considered; instead, it may be an accurate perception that too much dedication to work can carry some undesirable emotional side effects.

The multiple R for all predictors with this criterion was .49. There were no contradictions between the analyses represented in Tables 3 and 4.

Grades 4-9

The SET I instrument measures five aspects of teaching, as these older children see it. The first factor is Friendly and Cheerful. The part-correlates of this teaching measure, at these higher grade levels, were a lack of anxiety (ASD), a positive attitude toward authority (SRI) and a lack of stereotyped thinking (OWSC-Populars). It might be said that the children see a teacher's friendliness as the product of self-assurance, a comfortable relationship with authority figures, and also the ability to think for herself, not just deal in cliches.

The multiple R of all predictors with this criterion was .48, a little short of the .56 required for significance at the .05 level. There is a contradiction between the non-significant Pearson r of .11 and the almost-significant part-correlation of -.23, for OWSC-Populars; but this may simply be within the error-range of these correlations.

Knowledgeable, Poised teachers, as judged by their students, tended to express little anxiety (on the ASD); were much less likely to express strongly positive attitudes toward children (on the SRI); and tended to be conventional in their thinking (OWSC-Populars). The children seem to be impressed by the somewhat cool, firm self-assurance of such teachers, who perhaps are a little convention-bound and who are convinced of their own rightness.

The multiple R of all predictors with this variable was a significant .56. The data in Tables 3 and 4 correspond logically, for this criterion.

Lively, Interesting teachers, on the other hand, were those who expressed quite positive attitudes about people in authority (SRI) but who did not find work an unalloyed object for devotion (SRI); who were unevasively willing to speak their minds (OWSC-Evasions) and who were a little freer than average in voicing occasional hostility toward people of their own age or toward people in authority (OWSC-Hostility).

The multiple R of all predictors with this criterion was a non-signific. .46. There were no contradictory patterns in the data shown in Tables 3 and 4.

Teachers who were rated high on Firm Control tended to: be practical rather than idealistic (ASD); express positive feelings toward authority (SRI); give fewer than average stereotyped responses (OWSC-Populars); and were less likely to express feelings of anxiety or depression (OWSC). This pattern suggests a picture of reasoned but firm, businesslike purpose which the children simply recognize and respect for what it is.

The multiple R of all predictors with this variable was a significant .56. Comparing Tables 3 and 4, the significant Pearson r of .24 for the SRI-Hope score was washed out in the part-correlation analysis. Logically, it would appear likely to co-vary with the several measures of self-assurance that

contributed to Firm Control. The other discrepancy is the Pearson r of .01 and the part-correlation of -.25 for OWSC-Populars, though it is not a reversal.

Non-Directive teachers, as identified by their students, were extroverted (ASD-Introversion), idealistic (ASD-Idealism), but not wholly convinced of their personal attractiveness (ASD-Attractiveness). They expressed less positive feelings toward other people in general (SRI-Others) but highly positive attitudes toward authority (SRI-Authority). They were somewhat less positive toward work (SRI-Work) and somewhat less optimistic about the future (SRI-Hope) than teachers who were described as more directive. The pattern as a whole suggests that some instances of non-directive teaching may be more the result of teacher self-doubt and general uncertainty than the result of deliberate concern with giving children a chance to "learn how to learn" through their own exploratory efforts.

Indeed, such motivations may underly almost half of the instances of non-directive teaching, insofar as the multiple R for all predictors with this criterion was a substantial .65, accounting for 43% of the variance in this aspect of teaching. Such an interpretation would seem to be contradicted by the Pearson r of -.24 between OWSC-Anxiety and Non-Directiveness, but this negative relationship disappeared in the part-correlation analysis (see Table 4).

Predictions of Self-Ratings and Supervisor Ratings from the Personality Measures

Only two or three personality characteristics showed significant unique contributions to any of these three criterion measures. The overall Self-rating mean had a part correlation of -.18 with attitude toward parents (SRI); and -.21 with depressive feelings (OWSC).

On the other hand, when all the predictors in the battery were simultaneously related to this self-rating, the multiple R was .53, significant at the .01

level. A positive, generalized self-evaluation, partially represented on many of the personality scales, was positively related to self-appraisal as a teacher, as Table 3 demonstrates.

The mean of the ratings by the Cooperating Classroom Teacher had partial correlations of -.43 with self-described optimism and warmth (ASD-Attitude); -.27 with abrasiveness (ASD-Behavior); and -.20 with attitude toward own parents (SRI-Parents). What little definition the data provide suggests that supervising teachers prefer student teachers who are not aggressively optimistic or too self-assured (ASD-Attitude); who are pleasant rather than abrasive (ASD-Behavior); but who are emancipating themselves from uncritical ties to their parents (SRI-Parents).

The multiple R for all predictors with this supervising teacher rating was an almost-significant .42 (.45 would be significant at the .05 level). The most notable contrast between Tables 3 and 4 occur on this criterion. The twelve significant Pearson r's in Table 3 dwindle to two significant and one near-significant unique contributor in Table 4, to the variance in the criterion.

The mean of the TEF ratings by the College Supervisor had partial correlations only of -.32 with optimism and warmth (ASD-Attitude) and .17 with idealism (ASD). The multiple R for all predictors was .41, which approached but did not reach the .05 level of significance.

The predictive power of the personality measures was lower for the two adult ratings than for the student ratings. This may be due to the somewhat lower reliability of the supervisor's ratings, to their greater skewness, or to a genuine difference in the standards of judgement used by students and adults.

The direction of the relationships of personality characteristics to criteria was the same for almost all criteria. One exception was that being a little

depressed correlated positively with two of the pupil rating factors in the primary grades (Rapport, and Lack of Negativity) whereas it correlated negatively with Firm Control, as rated by the older students, and also negatively with TEF Self-rating.

Another exception involved the OWSC-Popular score. Interpreted here as representing somewhat stereotyped thinking (Veldman believes it may reflect a healthy normality of perception), it correlated negatively with SET II Rapport, Fosterance of Self-Esteem, and Lack of Negativity; negatively with SET I Friendly, and Firm Control; but positively with SET I Knowledgeable, Poised, and Non-Directive.

One other exception involved idealism (ASD), which correlated negatively with SET I Firm Control but positively with Non-Directiveness.

Otherwise, whenever a predictor characteristic was positively related to any one teaching criterion (i.e., the part correlation coefficients in Table 3), if it was related significantly to any other criterion, it was also in a positive direction. Conversely, a negative correlation with any one criterion was matched by negative correlations on any other criterion where the relationship reached significance. Since, with the possible exception of Non-Directive, all the criterion scales ran in the same direction, from less desirable to more desirable, the consistent direction of the correlations with specific personality characteristics suggests a certain degree of stability and meaningfulness in this relationship between personality and teaching behavior, partial though the relationships may be and imperfect as the measuring instruments certainly are.

Discussion

The general level and direction of the relationships between the personality measures and the teaching criteria were reasonably systematic and consistent,

although just near or at the most modest level of trustworthy significance. The three instruments, as a group, accounted for 17% to 24% of the variance in adult supervisor ratings (in all grade levels combined) and in the student ratings in the primary grades. In grades four through nine, the COMPASS scales accounted for 23% to 43% of the variance in student ratings, depending on the criterion at issue. Considering the quite transparent nature of the self-report ASD and SRI questionnaires and of the One-Word Sentence Completion, to some extent; considering the many, unidentified variations in the 141 classrooms where the student teachers performed; and considering the less than perfect reliability of all the measures; taking such factors into account, the degree of consistency in the findings suggests certain general conclusions:

1. The personality characteristics of teachers (student teachers, at least) do have a moderately predictable influence on the way their teaching is perceived and evaluated by both students and supervisors. The present data leave a majority of the variance in teaching still to be explained, undoubtedly by pupil differences and other situational factors, as well as by cognitive and behavioral attributes of the teachers which were not measured by the three instruments used in this study.

2. The level of validity of prediction is decidedly not strong enough to justify using these personality measures in a purely psychometric way, as the basis for administrative decisions such as admitting or excluding individuals from a teacher education program. Adding and testing other kinds of measures, of other attributes, might eventually lead to a battery which could be used administratively with a certain degree of rough justice; but it would almost certainly be rough.

3. On the other hand, the validity of these measures seems sufficient to infer that such data, in the hands of wise assessors, can be useful in determining personal characteristics of a given student teacher which are likely to bear on that teacher's future behavior in the classroom, and to forecast some of the probable effects on students. Read for content and pattern, more than for individual scores, such a battery can be useful for individual guidance and for mapping instructional strategies to suit the salient learning needs of a given student teacher. The findings thus lend support to the use of Personalized Assessment Feedback (Fuller, Bown, Newlove, Brown, 1972; Peck, Bown and Veldman, 1964) and subsequent, connected procedures in the Personalized Teacher Education Program. The full COMPASS battery used for that purpose includes several additional instruments, moreover, which amplify the data base for individualizing instruction.

4. The particular nature of the relationships found between personality and teaching hold some interest, too. Teachers who elect to work in the primary grades have somewhat different criteria to meet, to satisfy their students, than teachers at higher grade levels. This may be due to differences in student needs and values, to differences in the teachers who elect one or the other grade level, or a combination of both. It is known, from the developmental work on the SET I and II instruments, that primary children and older children use somewhat different sets of conceptual dimensions when they evaluate their teachers.

The SET II began with substantially the same item-pool as the SET I, with only a simplification of wording which seldom seemed to change the meaning of an item. The four factors that emerged empirically from the SET II, however, are a somewhat different way of identifying what the children see as critical elements of teaching, as compared with the five factors which were empirically generated by older children, using the SET I items. It is known, too, from a

good deal of research that teachers who prefer to work at the primary level usually show certain values and relational behavior which differs from the personal value-priorities and relational behavior of teachers who prefer to work with older children.

In the present study, the primary grade children seem to prefer teachers who are friendly (low OWSC-Hostility) but who do not paint themselves as rosily optimistic and warmly goodnatured to a (perhaps unrealistically) high degree (ASD-Attitude). The children prefer teachers who are outgoing (ASD-Introversion), who are emancipated from overly close ties to their own parents (an inference from the SRI-Parents pattern), and who think for themselves (OWSC-Populars). They like teachers, too, who show a mild amount of anxiety or depression -- probably not in the literal sense, but perhaps because such teachers are apt to be more self-critical than to be punitively critical toward their pupils. Clearly, the positive emotional quality of teacher-student relationships is a central concern of these young children. This is apparent in the very factors these children have identified as their major "yardsticks" for evaluating teachers (the SET II factors). It also seems apparent in the particular personality attributes that correlate with the SET ratings.

The older children, somewhat by way of contrast, prefer non-anxious teachers (ASD-Anxiety) who are practical-minded (ASD-Idealism), who respect and feel comfortable with people in authority (SRI), yet who are self-assured enough to be able to voice annoyance with other adults (OWSC-Hostility) if they feel it is justified. There is more complexity to the evaluative system of these children, however. They tend to regard conventionality of thought in a teacher (OWSC-Populars) as a good thing insofar as it relates to qualities of Friendly

Cheerfulness, Poised Knowledgeability, and Non-Directiveness; but they see it as a slight deterrent to Firm Control. To the degree that they weigh the presence of an anxious or depressive tendency in a teacher, they see it only as a deterrent to Firm Control (OMSC correlations) or to Cheerfulness or Poise (ASD-Anxiety correlations).

Of particular interest is the quite strong, negative relationship between their evaluation of a Teacher's Poise and that teacher's avowed liking for children (SRI-Children). This does not mean that they prefer teachers who actively dislike children; rather, they prefer teachers who do not go overboard in exuding a "love for children," to judge from the score distributions on the SRI-Children scale.

The older children, as the SET I factors indicate, put somewhat less stress than primary children on being treated in a kindly, confidence-enhancing way. They add the new dimensions of Poised Knowledgeability and Firm Control as important things for which they respect teachers. They appear to regard Non-Directive teaching behavior in a somewhat equivocal way, perhaps about the same way adults do: favoring it, when it is a perceptive way of encouraging student initiative, but disfavoring it when it seems to grow out of a teacher's lack of self-assurance or self-direction. Their "yardstick" dimensions reconfirm the findings of Ryans (1960).

5. While the adult criterion ratings showed somewhat lower reliability, probably as a function of the very small number of judges, they correlated (by multiple regression) with the COMPASS battery almost to the level of the SET II relationship, though less than the SET I level. It was not possible to identify many specific teacher attributes that accounted for these adult

evaluations, and therefore the nature of the very considerable differences between the adults' and the students' evaluative frameworks cannot be identified and specified. Some G-factor or halo-factor may have operated in the adult ratings, so that their multiple R with the COMPASS battery reached near-significance, but few elements of that battery stood out as strong, individually identifiable predictors. This tends to support the practical utility of the COMPASS battery, but it contributes little to an understanding of the concepts and "yardsticks" the adult raters use in evaluating teachers, insofar as their ratings are influenced by the kinds of attributes measured by the three personality instruments.

In sum, this study found a consistent pattern of relationships between certain personality characteristics of student teachers and the way they were seen and evaluated by their pupils and by their supervisors. New research should test these relationships in new samples, perhaps especially with populations of disadvantaged and minority children, who were not included in this study. To raise the multiple R's above the present level probably will require identifying additional teacher characteristics, of both cognitive and affective kinds. It may also require identifying new measures of context-factors (pupil characteristics, school value system, particular supervisor characteristics, etc.) which influence the impact it is possible for a student teacher to make. Finally, it appears extremely important to tie all performance ratings far more specifically and firmly to relevant, adequately complete, direct measures of the actual consequences of different kinds of teaching behavior on the cognitive and affective learning of different kinds of children.

References

Bown, O. H. and Veldman, D. J. Scoring procedure and college freshman norms for the Self-Report Inventory. Research and Development Center for Teacher Education, The University of Texas at Austin, 1967.

Fuller, F. F., Bown, O. H., Newlove, B. and Brown, G. G. Personal assessment feedback counseling for teachers: A manual for counseling psychologists. Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Haak, R., Kleiber, D. and Peck, R. F. Student evaluation of teacher instrument II manual. The Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Peck, R. F. Report on the feasibility of programming an electronic computer for the psychological analysis of individuals. MHTE Research Memo No. 8, The University of Texas, May, 1960.

Peck, R. F. Personalized education: An attainable goal in the seventies. In Needs of Elementary and Secondary Education for the Seventies, a Compendium of Policy Papers, compiled by the General Subcommittee on Education of the Committee on Education and Labor, House of Representatives, Ninety-First Congress, First Session. U. S. Government Printing Office, Washington: March, 1970, 643-649.

Peck, R. F. et al. Personalized teacher education (Basic Program Plan submitted to U. S. Office of Education, Contract No. OE-6-10-108.) Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Peck, R. F., Bown, O. H. and Veldman, D. J. Mental health in teacher education. Journal of Teacher Education, 1964, 15, 3, 319-327.

Peck, R. F. and Veldman, D. J. Computer analysis of sentence completion data. American Psychologist, 1966, 21, p. 678.

Ryans, D. G. Characteristics of Teachers, Their Description, Comparison and Appraisal. Washington, D. C.: American Council on Education, 1960.

Veldman, D. J. Pupil evaluation of student teachers and their supervisors. Journal of Teacher Education, 1970, 21, 165-167.

Veldman, D. J. Automated sentence completion screening. Research Methodology Monograph No. 12, Research and Development Center for Teacher Education, The University of Texas at Austin, Summer, 1971.

Veldman, D. J. and Parker, G.V.C. Adjective rating scales for self-description. Multivariate Behavioral Research, 1970, 5, 295-302.

Veldman, D. J. and Peck, R. F. The influence of teacher and pupil sex on pupil evaluations of student teachers. Journal of Teacher Education, December 1964, Vol. 15, No. 4, 393-396.

TABLE 1

Reliability of SET I and SET II

A. Inter-Class Reliability of SET I

Mean reliability coefficients for the five factors, across 50 teachers, each rated by two classes of pupils: Factor 1: .92; 2: .72; 3: .91; 4: .81; 5: .89.

B. Test-Retest Reliability of SET II

(rated ten days apart, by 1,040 pupils)

Level	Range of correlations	Average reliability
kindergarten	.63 - .90	.74
grade 1	.67 - .92	.82
grades 2 and 3	.68 - .89	.80
grades 4, 5, 6 (Mex-Am pupils)	.67 - .94	.80

TABLE 2

Range of Correlations of Adult and Pupil Evaluations

	SET II (4 factors)	SET I (5 factors)
cooperating teacher	-.31 to .29	-.01 to .14
supervisor	-.20 to .23	.03 to .16
self	-.06 to .12	-.02 to .16

Table 3

Pearson Correlations of Personality Variables with Teaching Performance Criteria

COMPASS MEASURE	SET II				SET I					TEF Ratings		
	Stimulating Rapport	Effective Teaching	Fosters Pupil Self-Esteem	Lack of Negativity	Friendly, Cheerful	Knowledgeable, Poised	Lively, Interesting	Firm Control	Non-Directive	Self	Cooperating Teacher	Supervising Instructor
ASD												
1. Attitude	-13	-10	03	-01	-02	14	12	01	01	.25	-04	-06
2. Behavior (Abrasive)	15	09	-02	06	13	-10	07	-22	05	-.25	-.20	-13
3. Efficiency	-12	-10	-02	-05	-09	08	-04	22	01	25	13	13
4. Introversion	-05	-11	-.26	-04	00	-04	-10	-00	-18	-.18	-07	-12
5. Anxiety	12	.23	-04	13	-.25	-14	-04	-03	-04	-05	-01	-11
6. Idealism	-03	16	-05	-04	-06	-06	04	-.28	22	.16	03	10
7. Attractive	-02	-06	15	02	-11	13	03	08	-05	.27	08	01
MAI												
1. Self	-08	-17	16	-11	08	-05	13	13	17	.29	15	11
2. Others	-07	00	07	-09	06	16	10	04	-11	.19	10	-03
3. Children	01	-07	07	-01	00	-18	05	06	04	.20	10	04
4. Authority	-01	-02	09	-06	19	20	13	14	19	.18	11	-01
5. Work	-14	-17	-06	-21	-10	05	-17	18	-07	12	11	07
7. Parents	-22	-13	-01	-17	14	09	08	16	-13	01	-08	-08
8. Hope	-11	-02	11	-09	00	13	-01	.24	-10	.23	09	07
OWSC												
1. Response Length	10	07	-11	-07	-06	-03	-13	13	-12	03	02	06
2. Repetitions	-20	-06	03	-10	-08	-01	05	10	-02	07	02	-02
3. Populars	-13	-20	03	-13	11	.24	08	01	15	08	11	03
5. Hostility	-09	13	-12	-01	-02	00	05	06	06	-05	-09	-08
6. Anxiety	22	.25	05	20	-15	-09	-18	-11	-.24	05	04	-06
7. Depression	17	18	-05	16	-13	-10	-02	-.35	09	-.26	-11	-11
Multiple r	49	46	45	49	48	56	46	56	66	53	42	41

.30 significant < .01
.23 significant < .05
r_M 56 Significant < .05

.21 significant < .01
.16 significant < .05
r_M 45 significant < .05

Table 4

Part Correlations of Variables Contributing a Significant Amount to Criterion Variance

Personality Measures	SET II				SET I					TEF Ratings		
ASD	Stimulating Rapport	Effective Teaching	Fosters Pupil Self-Esteem	Lack of Negativity	Friendly, Cheerful	Knowledgeable Poised	Lively, Interesting	Firm Control	Non-Directive	Self	Cooperating Teacher	Supervising Instructor
1. Attitude	(-26)	-32									-43	-32
2. Behavior (Abrasive)											-27	
4. Introversion			-28						-29			
5. Anxiety					-25	(-18)						
6. Idealism								-36	31			(17)
7. Attractive									-22			
SRI												
1. Self			(23)									
2. Others									-35			
3. Children						-54						
4. Authority					(22)		(22)	(18)	49			
5. Work				-31			-32		(-20)			
7. Parents	-27	(-17)	(-19)	-32						(-18)	(-20)	
8. Hope									-25			
OWSC												
1. Response Length			(-23)	-24								
2. Repetitions												
3. Populars		-31	-27	-24	(-23)	32		-25	30			
4. Evasion							(-19)					
5. Hostility	-34		-21	-22			(21)					
6. Anxiety	(16)	21						-22				
7. Depression	28			22				-34		-21		
Total Unique Variance	19	12	13	19	07	15	08	17	34	05	13	04
Total "Explained" Variance	24	21	20	24	23	31	21	31	43	28	18	17

(r) = significant contribution to criterion variance $<$ .10 probability
r = significant contribution to criterion variance $<$ .05 probability